

SMTP
Internet-Draft
Intended status: Standards Track
Expires: May 4, 2009

D. Crocker
Brandenburg InternetWorking
October 31, 2008

Internet Mail Architecture
draft-crocker-email-arch-11

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on May 4, 2009.

Abstract

Over its thirty-five year history, Internet Mail has changed significantly in scale and complexity, as it has become a global infrastructure service. These changes have been evolutionary, rather than revolutionary, reflecting a strong desire to preserve both its installed base and its usefulness. To collaborate productively on this large and complex system, all participants must work from a common view of it and use a common language to describe its components and the interactions among them. But the many differences in perspective currently make it difficult to know exactly what another participant means. To serve as the necessary common frame of reference, this document describes the enhanced Internet Mail architecture, reflecting the current service.

Internet-Draft

EMail Architecture

October 2008

Table of Contents

1.	Introduction	3
1.1.	History	3
1.2.	Document Conventions	6
2.	Responsible Actor Roles	7
2.1.	User Actors	7
2.2.	Mail Handling Service (MHS) Actors	10
2.3.	Administrative Actors	13
3.	Identities	16
3.1.	Mailbox	16
3.2.	Scope of Email Address Use	17
3.3.	Domain Names	17
3.4.	Message Identifier	18
4.	Services and Standards	19
4.1.	Message Data	22
4.2.	User-Level Services	27
4.3.	MHS-Level Services	29
4.4.	Transition Modes	33
4.5.	Implementation and Operation	33
5.	Mediators	34
5.1.	Alias	35
5.2.	ReSender	36
5.3.	Mailing Lists	37
5.4.	Gateways	39
5.5.	Boundary Filter	40
6.	Considerations	40
6.1.	Security Considerations	40
6.2.	IANA Considerations	41
6.3.	Internationalization	41
7.	References	42
7.1.	Normative	42
7.2.	Informative	44
Appendix A.	Acknowledgements	45
	Index	46
	Author's Address	48
	Intellectual Property and Copyright Statements	49

[1.](#) Introduction

Over its thirty-five year history, Internet Mail has changed significantly in scale and complexity, as it has become a global infrastructure service. These changes have been evolutionary, rather than revolutionary, reflecting a strong desire to preserve both its installed base and its usefulness. Today, Internet Mail is distinguished by many independent operators, many different components for providing service to users, as well as many different components that transfer messages.

Public collaboration on technical, operations, and policy activities of email, including those that respond to the challenges of email abuse, has brought a much wider range of participants into the technical community. To collaborate productively on this large and complex system, all participants must work from a common view of it and use a common language to describe its components and the interactions among them. But the many differences in perspective currently make it difficult to know exactly what another participant means.

It is the need to resolve these differences that motivates this document, which describes the realities of the current system. Internet Mail is the subject of ongoing technical, operations, and policy work, and the discussions often are hindered by different models of email service design and different meanings for the same terms.

To serve as the necessary common frame of reference, this document describes the enhanced Internet Mail architecture, reflecting the current service. The document focuses on:

- * Capturing refinements to the email model

- * Clarifying functional roles for the architectural components
- * Clarifying identity-related issues, across the email service
- * Defining terminology for architectural components and their interactions

1.1. History

The first standardized architecture for networked email specified a simple split between the user world, in the form of Mail User Agents (MUA), and the transfer world, in the form of the Mail Handling

Crocker

Expires May 4, 2009

[Page 3]

Internet-Draft

EMail Architecture

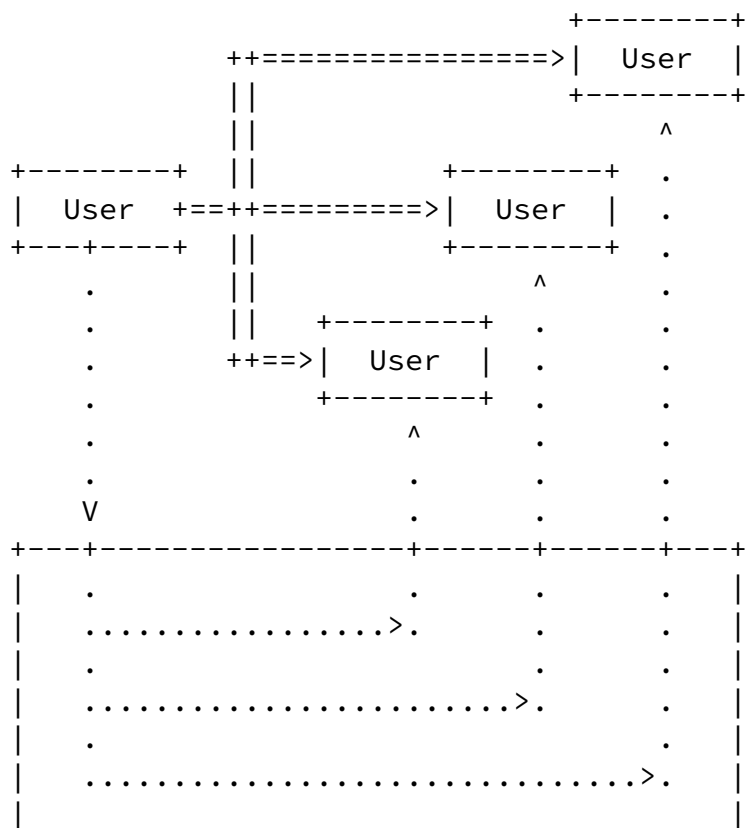
October 2008

Service (MHS), which is composed of Mail Transfer Agents (MTA). The MHS accepts a message from one User and delivers it to one or more other users, creating a virtual MUA-to-MUA exchange environment.

As shown in Figure 1, this defines two logical layers of interoperability. One is directly between Users. The other is among the components along the transfer path. In addition, there is interoperability between the layers, first when a message is posted from the User to the MHS and later when it is delivered from the MHS to the User.

The operational service has evolved, although core aspects of the service, such as mailbox addressing and message format style, remaining remarkably constant. The original distinction between the user level and transfer level remains, but with elaborations in each. The term "Internet Mail" is used to refer to the entire collection of user and transfer components and services.

For Internet Mail, the term "end-to-end" usually refers to a single posting and the set of deliveries that result from a single transit of the MHS. A common exception is group dialogue that is mediated, through a Mailing List; in this case, two postings occur before intended Recipients receive an Author's message, as discussed in [Section 2.1.3](#). In fact, some uses of email consider the entire email service, including Author and Recipient, as a subordinate component. For these services, "end-to-end" refers to points outside the email service. Examples are voicemail over email "[[RFC3801](#)]", EDI over email [[RFC1767](#)] and facsimile over email [[RFC4142](#)].



| Mail Handling Service (MHS) |
+-----+

Figure 1: Basic Internet Mail Service Model

End-to-end Internet Mail exchange is accomplished by using a standardized infrastructure with these components and characteristics:

- * An email object
- * Global addressing
- * An asynchronous sequence of point-to-point transfer mechanisms
- * No prior arrangement between MTAs or between Authors and Recipients
- * No prior arrangement between point-to-point transfer services over the open Internet
- * No requirement for Author, Originator, or Recipients to be online at the same time

The end-to-end portion of the service is the email object, called a "message." Broadly, the message itself distinguishes control information, for handling, from the Author's content.

A precept to the design of mail over the open Internet is permitting user-to-user and MTA-to-MTA interoperability without prior, direct arrangement between the independent administrative authorities responsible for handling a message. All participants rely on having the core services universally supported and accessible, either directly or through Gateways that act as translators between Internet Mail and email environments conforming to other standards. Given the importance of spontaneity and serendipity in interpersonal communications, not requiring such prearrangement between participants is a core benefit of Internet Mail and remains a core requirement for it.

Within localized networks at the edge of the public Internet, prior administrative arrangement often is required and can include access control, routing constraints, and configuration of the information query service. Although recipient authentication has usually been required for message access since the beginning of Internet Mail, in recent years it also has been required for message submission. In these cases, a server validates the client's identity, whether by explicit security protocols or by implicit infrastructure queries to identify "local" participants.

1.2. Document Conventions

References to structured fields of a message use a two-part dotted notation. The first part cites the document that contains the specification for the field and the second is the name of the field. Hence <[RFC2822](#).From> is the From: header field in an email content header and <[RFC2821](#).MailFrom> is the address in the SMTP "Mail From" command.

When occurring without the [RFC2822](#) qualifier, header field names are shown with a colon suffix. For example, From:.

References to labels for actors, functions or components have the first letter capitalized.

Also, the key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#) [[RFC2119](#)] [[RFC2119](#)].

RFC EDITOR: Remove the following paragraph before publication.

Discussion venue: Please direct discussion about this document to the IETF-SMTP mailing list <<http://www.imc.org/ietf-smtp>>.

[2.](#) Responsible Actor Roles

Internet Mail is a highly distributed service, with a variety of actors playing different roles. These actors fall into three basic types:

- * User
- * Mail Handling Service (MHS)
- * ADministrative Management Domain (ADMD)

Although related to a technical architecture, the focus on actors concerns participant responsibilities, rather than functionality of modules. For that reason, the labels used are different from those used in classic email architecture diagrams.

[2.1.](#) User Actors

Users are the sources and sinks of messages. Users can be people, organizations, or processes. They can have an exchange that iterates, and they can expand or contract the set of users that participate in a set of exchanges. In Internet Mail, there are four types of Users:

- * Authors
- * Recipients
- * Return Handlers
- * Mediators

Figure 2 shows the primary and secondary flows of messages among them.

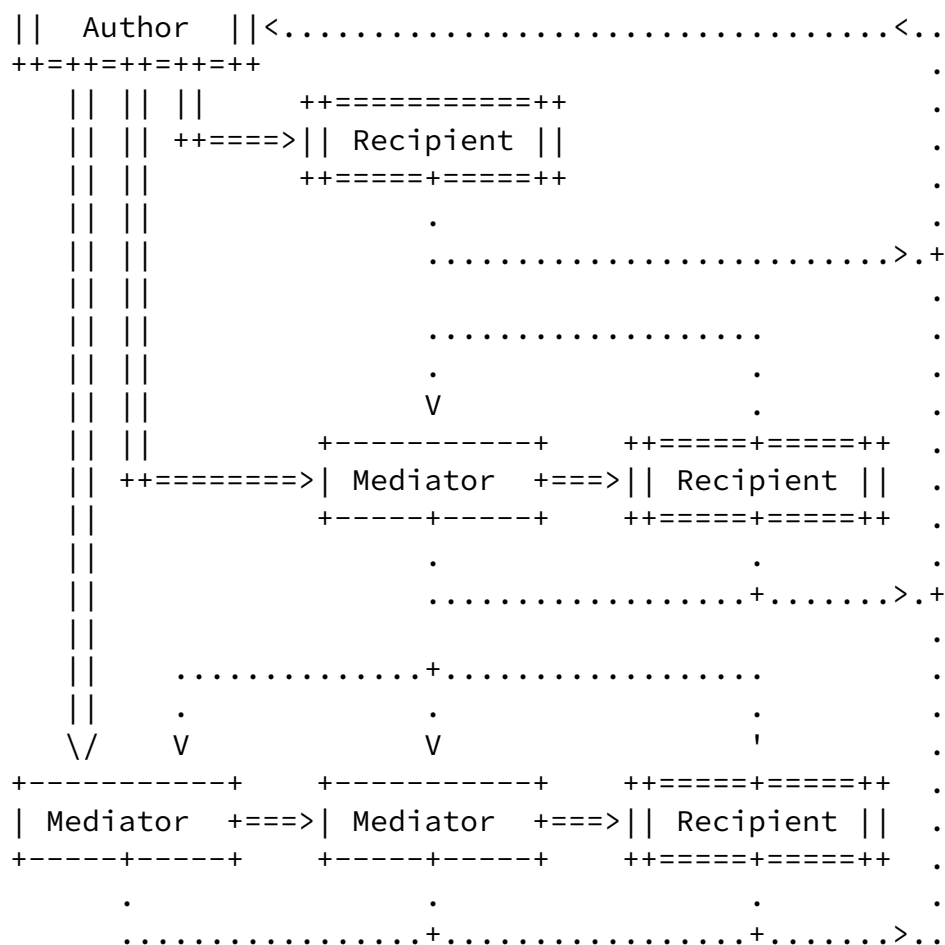


Figure 2: Relationships Among User Actors

From the user perspective, all mail transfer activities are performed by a monolithic Mail Handling Service (MHS), even though the actual service can be provided by many independent organizations. Users are customers of this unified service.

Whenever any MHS actor sends information to back to an Author or Originator in the sequence of handling a message, that actor is a User.

[2.1.1.](#) Author

The Author is responsible for creating the message, its contents, and its list of recipient addresses. The MHS transfers the message from the Author and delivers it to the Recipients. The MHS has an Originator role ([Section 2.2.1](#)) that correlates with the Author role.

[2.1.2.](#) Recipient

The Recipient is a consumer of the delivered message. The MHS has a Receiver role ([Section 2.2.4](#)) correlates with the Recipient role. This is labeled Recv in Figure 3.

Any Recipient can close the user communication loop by creating and submitting a new message that replies to the Author. An example of an automated form of reply is the Message Disposition Notification (MDN), which informs the Author about the Recipient's handling of the message. (See [Section 4.1.](#))

The Return Handler, also called "Bounce Handler," receives and services notifications that the MHS generates, as it transfers or delivers the message. These notices can be about failures or completions and are sent to an address that is specified by the Originator<<initial def>> . This Return handling address (also known as a Return address) might have no visible characteristics in common with the address of the Author or Originator.

[2.1.3.](#) Mediator

A Mediator receives, aggregates, reformulates, and redistributes messages among Authors and Recipients who are the principals in protracted exchanges. This activity is easily confused with the underlying MHS transfer exchanges. However, each serves very different purposes and operates in very different ways.

When mail is delivered to the Mediator specified in the [RFC2821](#).RcptTo command, the MHS handles it the same way as for any other Recipient. The MHS sees each posting and delivery activity between sources and sinks as independent; it does not see subsequent re-posting as a continuation of a process. Because the Mediator originates messages, it can receive replies. Hence, when submitting messages, the Mediator is an Author. So a Mediator really is a full-fledged User. Mediators are considered extensively in [Section 5](#).

The distinctive aspects of a Mediator are outside the MHS. A Mediator preserves the Author information of the message it reformulates and is permitted to make meaningful changes to the message content or envelope. The MHS sees a new message, but users receive a message that they interpret as being from, or at least initiated by, the Author of the original message. The role of a Mediator is not limited to merely connecting other participants; the Mediator is responsible for the new message.

A Mediator's role is complex and contingent, for example, modifying and adding content or regulating which users are allowed to

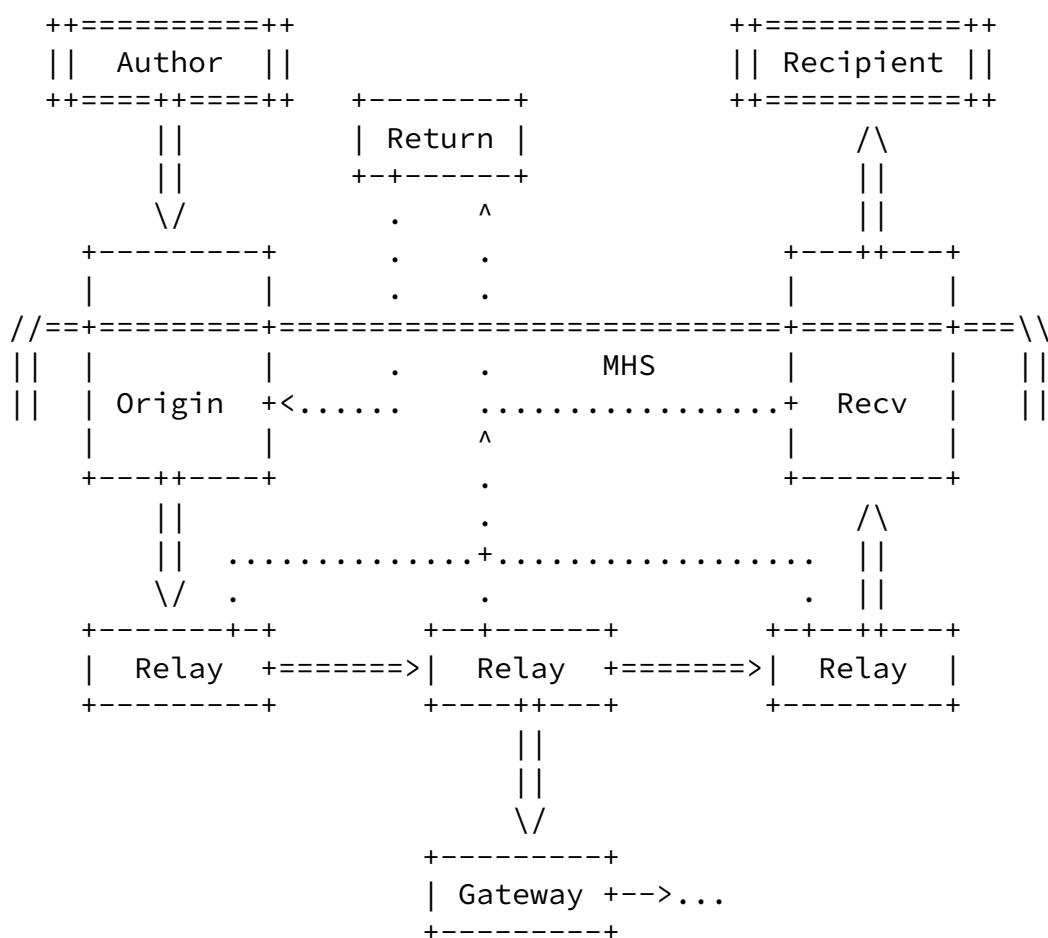
participate and when. The common example of this role is a group Mailing List. In a more complex use, a sequence of Mediators could perform a sequence of formal steps, such as reviewing, modifying, and approving a purchase request.

A Gateway is a particularly interesting form of Mediator. It is a hybrid of User and Relay that connects heterogeneous mail services. Its purpose is to emulate a Relay. For a detailed discussion, see [Section 2.2.3](#).

[2.2](#). Mail Handling Service (MHS) Actors

The Mail Handling Service (MHS) performs a single end-to-end transfer on behalf of the Author to reach the Recipient addresses specified in the original [RFC2821](#).RcptTo commands. Exchanges that are either mediated or iterative and protracted, such as those used for collaboration over time are handled by the User actors, not by the MHS actors.

Figure 3 shows the relationships among transfer participants in Internet Mail. Although it shows the Originator (labeled Origin) as distinct from the Author and Receiver (labeled Recv) as distinct from Recipient, each pair of roles usually has the same actor. Transfers typically entail one or more Relays. However direct delivery from the Originator to Receiver is possible. Intra-organization mail services usually have only one Relay.



2.2.1. Originator

The Originator ensures that a message is valid for posting and then submits it to a Relay. A message is valid if it conforms to both

Internet Mail standards and local operational policies. The Originator can simply review the message for conformance and reject it if it finds errors, or it can create some or all of the necessary information. In effect, the Originator is responsible for the functions of the Mail Submission Agent.

The Originator operates with dual allegiance. It serves the Author and can be the same entity. But its role in assuring validity means that it MUST also represent the local operator of the MHS, that is, the local ADministrative Management Domain (ADMD).

The Originator also performs any post-submission, Author-related administrative tasks associated with message transfer and delivery. Notably, these tasks pertain to sending error and delivery notices, enforcing local policies, and dealing with messages from the Author that prove to be problematic for the Internet. The Originator is

Crocker

Expires May 4, 2009

[Page 11]

Internet-Draft

EMail Architecture

October 2008

accountable for the message content, even when it is not responsible for it. The Author creates the message, but the Originator handles any transmission issues with it.

[2.2.2.](#) Relay

The Relay performs MHS-level transfer-service routing and store-and-forward, by transmitting or retransmitting the message to its Recipients. The Relay adds trace information [[RFC2505](#)] but does not modify the envelope information or the message content semantics. It can modify message content representation, such as changing the form of transfer encoding from binary to text, but only as required to meet the capabilities of the next hop in the MHS.

A Mail Handling Service (MHS) network consists of a set of Relays. This network is above any underlying packet-switching network that might be used and below any Gateways or other Mediators.

In other words, email scenarios can involve three distinct architectural layers, each providing its own type of data of store-and-forward service:

- * User Mediators

- * MHS Relays
- * Packet Switches

The bottom layer is the Internet's IP service. The most basic email scenarios involve Relays and Switches.

Aborting a message transfer makes the Relay an Author because it must send an error message to the Return address. The potential for looping is avoided by omitting a Return address from this message.

[2.2.3.](#) Gateway

A Gateway is a hybrid of User and Relay that connects heterogeneous mail services. Its purpose is to emulate a Relay and the closer it comes to this, the better. A Gateway operates as a User when it needs the ability to modify message content.

Differences between mail services can be as small as minor syntax variations, but they usually encompass significant, semantic distinctions. One difference could be email addresses that are hierarchical and machine-specific rather than a flat, global

namespace. Another difference could be support for text-only content or multi-media. Hence the Relay function in a Gateway presents a significant design challenge, if the resulting performance is to be seen as nearly seamless. The challenge is to ensure user-to-user functionality between the services, despite differences in their syntax and semantics.

The basic test of Gateway design is whether an Author on one side of a Gateway can send a useful message to a Recipient on the other side, without requiring changes to any components in the Author's or Recipient's mail services other than adding the Gateway. To each of these otherwise independent services, the Gateway appears to be a native participant. But the ultimate test of Gateway design is whether the Author and Recipient can sustain a dialogue. In particular, can a Recipient's MUA automatically formulate a valid Reply that will reach the Author?

[2.2.4.](#) Receiver

The Receiver performs final delivery or sends the message to an alternate address. It can also perform filtering and other policy enforcement immediately before or after delivery.

2.3. Administrative Actors

Administrative actors can be associated with different organizations, each with its own administrative authority. This operational independence, coupled with the need for interaction between groups, provides the motivation to distinguish among Administrative Management Domains (ADMDs). Each ADMD can have vastly different operating policies and trust-based decision-making. One obvious example is the distinction between mail that is exchanged within an organization and mail that is exchanged between independent organizations. The rules for handling both types of traffic tend to be quite different. That difference requires defining the boundaries of each, and this requires the ADMD construct.

Operation of Internet Mail services is carried out by different providers (or operators). Each can be an independent ADMD. This independence of administrative decision-making defines boundaries that distinguish different portions of the Internet Mail service. A department that operates a local Relay, an IT department that operates an enterprise Relay, and an ISP that operates a public shared email service can be configured into many combinations of administrative and operational relationships. Each is a distinct ADMD, potentially having a complex arrangement of functional components. Figure 4 depicts relationships among ADMDs. The benefit of the ADMD construct is to facilitate discussion about designs,

policies and operations that need to distinguish between internal issues and external ones.

The architectural impact of the need for boundaries between ADMDs is discussed in [[Tussle](#)]. Most significant is that the entities communicating across ADMD boundaries typically have the added burden of enforcing organizational policies concerning external communications. At a more mundane level, routing mail between ADMDs can be an issue, such as needing to route mail for partners over specially trusted paths.

These are the basic types of ADMDs:

Edge: Independent transfer services in networks at the edge of the open Internet Mail service.

Consumer: This might be a type of Edge service, as is common for web-based email access.

Transit: Mail Service Providers (MSP) that offer value-added capabilities for Edge ADMDs, such as aggregation and filtering.

The mail-level transit service is different from packet-level switching. End-to-end packet transfers usually go through intermediate routers; email exchange across the open Internet can be directly between the Boundary MTAs of Edge ADMDs. This distinction between direct and indirect interaction highlights the differences discussed in [Section 2.2.2](#)

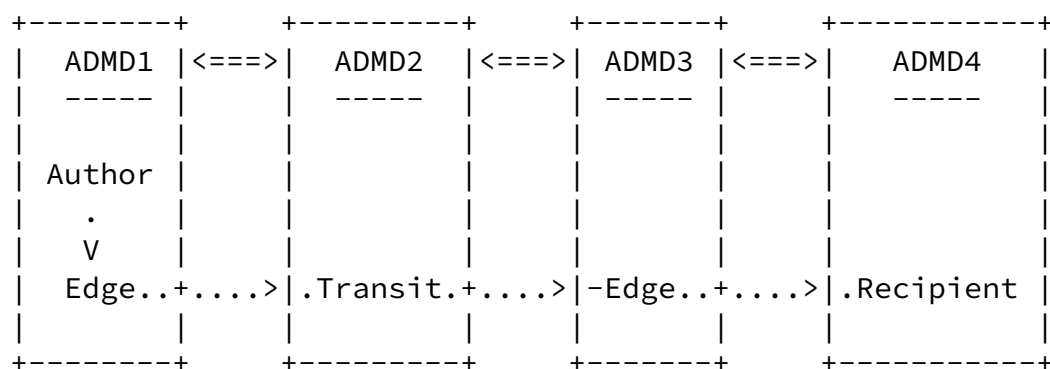


Figure 4: Administrative Domain (ADMD) Example

Edge networks can use proprietary email standards internally. However the distinction between Transit network and Edge network transfer services is significant because it highlights the need for concern over interaction and protection between independent administrations. In particular, this distinction calls for

additional care in assessing the transitions of responsibility and the accountability and authorization relationships among participants in message transfer.

The interactions of ADMD components are subject to the policies of that domain, which cover concerns such as these:

- * Reliability
- * Access control
- * Accountability
- * Content evaluation and modification

These policies can be implemented in different functional components, according to the needs of the ADMD. For example, see [[RFC5068](#)].

Consumer, Edge, and Transit services can be offered by providers that operate component services or sets of services. Further, it is possible for one ADMD to host services for other ADMDs.

These are common examples of ADMDs:

Enterprise Service Providers:

These ADMDs operate the internal data and/or the mail services within an organization.

Internet Service Providers (ISP):

These ADMDs operate the underlying data communication services, which are used by one or more Relay and User. ISPs are not responsible for performing email functions, but they can provide an environment in which those functions can be performed.

Mail Service Providers:

These ADMDs operate email services, such as for consumers or client companies.

Practical operational concerns demand that providers be involved in administration and enforcement issues. This involvement can extend to operators of lower-level packet services.

[3.](#) Identities

Internet Mail uses three forms of identity: mailbox, domain name, and message-ID. Each must be globally unique.

[3.1.](#) Mailbox

"A mailbox sends and receives mail. It is a conceptual entity which does not necessarily pertain to file storage." [[RFC2822](#)]

A mailbox is specified as an Internet Mail address <addr-spec>. It has two distinct parts, separated by an at-sign (@). The right side is a globally interpreted domain name associated with an ADMD. Domain names are discussed in [Section 3.3](#). Formal Internet Mail addressing syntax can support source routes, to indicate the path through which a message ought to be sent. The use of source routes is not common and has been deprecated in [[RFC2821](#)].

The portion to the left of the at-sign contains a string that is globally opaque and is called the <local-part>. It is to be interpreted only by the entity specified by the address's domain name. Except as noted later in this section all other entities MUST treat the <local-part> as an uninterpreted literal string and MUST preserve all of its original details. As such its public distribution is equivalent to sending a Web browser "cookie" that is only interpreted upon being returned to its creator.

Some local-part values have been standardized, for contacting personnel at an organization. These names cover common operations and business functions. [[RFC2142](#)]

It is common for sites to have local structuring conventions for the left-hand side <local-part> of an <addr-spec>. This permits sub-addressing, such as for distinguishing different discussion groups used by the same participant. However it is worth stressing that these conventions are strictly private to the user's organization and MUST NOT be interpreted by any domain except the one listed in the right side of the <addr-spec>. The exceptions are those specialized services that conform to public, standardized conventions, as noted

below.

A few types of addresses elaborate on basic email addressing, with a standardized, global schema for the <local-part>. Include are conventions between authoring systems and Gateways. They are invisible to the public email transfer infrastructure. When an Author is explicitly sending through a Gateway out of the Internet, coding conventions for the <local-part> allow the Author to formulate instructions for the Gateway. Standardized examples of such conventions are the telephone numbering formats for VPIM [[RFC3801](#)], such as:

+16137637582@vpim.example.com

and iFax [[RFC3192](#)], such as:

FAX=+12027653000/T33S=1387@ifax.example.com

[3.2.](#) Scope of Email Address Use

Email addresses are being used far beyond their original role in email transfer and delivery. In practical terms, an email address string has become the common identifier for representing online identity. Hence, it is essential to be clear about both the nature and role of an identity string in a particular context and the entity responsible for setting that string. For example, see [Section 4.1.4](#), [Section 4.3.3](#) and [Section 5](#).

[3.3.](#) Domain Names

A domain name is a global reference to an Internet resource, such as a host, a service, or a network. A domain name usually maps to one or more IP Addresses. Conceptually, the name can encompass an organization, a collection of machines integrated into a homogeneous service, or a single machine. A domain name can be administered to refer to individual users, but this is not common practice. The name is structured as a hierarchical sequence of names, separated by dots (.), with the top of the hierarchy being on the right end of the sequence. Domain names are defined and operated through the Domain Name System (DNS) [[RFC1034](#)], [[RFC1035](#)], [[RFC2181](#)].

When not part of a mailbox address, a domain name is used in Internet Mail to refer to the ADMD or to the host that took action upon the message, such as providing the administrative scope for a message identifier or performing transfer processing.

[3.4.](#) Message Identifier

There are two standardized tags for identifying messages: Message-ID: and ENVID. A Message-ID: pertains to content, and an ENVID pertains to transfer.

[3.4.1.](#) Message-ID

Internet Mail standards provide for, at most, a single Message-ID:. The Message-ID: for a single message, which is a user-level tag, has a variety of uses including threading, aiding identification of duplicates, and DSN tracking. [[RFC2822](#)]. The Originator assigns the Message-ID:. The Recipient's ADMD is the intended consumer of the Message-ID:, although any actor along the transfer path can use it.

Message-ID: MUST be globally unique. Its format is similar to that of a mailbox, with two distinct parts, separated by an at-sign (@). Typically, the right side specifies the ADMD or host that assigns the identifier, and the left side contains a string that is globally opaque and serves to uniquely identify the message within the domain referenced on the right side. The duration of uniqueness for the message identifier is undefined.

When a message is revised in any way, the decision whether to assign a new Message-ID: requires a subjective assessment to determine whether the editorial content has been changed enough to constitute a new message. [[RFC2822](#)] states that "a message identifier pertains to exactly one instantiation of a particular message; subsequent revisions to the message each receive new message identifiers." Yet experience suggests that some flexibility is needed. An impossible test is whether the recipient will consider the new message to be equivalent to the old one. For most components of Internet Mail, there is no way to predict a specific recipient's preferences on this

matter. Both creating and failing to create a new Message-ID: have their downsides.

Here are some guidelines and examples:

- * If a message is changed only in form, such as character-encoding, it is still the same message.
- * If a message has minor additions to the content, such as a mailing list tag at the beginning of the [RFC2822](#).Subject header field, or some mailing list administrative information added to the end of the primary body-part text, it is probably the same message.

Crocker

Expires May 4, 2009

[Page 18]

Internet-Draft

EMail Architecture

October 2008

- * If a message has viruses deleted from it, it is probably the same message.
- * If a message has offensive words deleted from it, some recipients will consider it the same message, but some will not.
- * If a message is translated into a different language, some recipients will consider it the same message, but some will not.
- * If a message is included in a digest of messages, the digest constitutes a new message.
- * If a message is forwarded by a recipient, what is forwarded is a new message.
- * If a message is "redirected", such as using [RFC2822](#) "Resent-*" header fields, some recipients will consider it the same message, but some will not.

The absence of both objective, precise criteria for re-generating a Message-ID: and strong protection associated with the string means that the presence of an ID can permit an assessment that is marginally better than a heuristic, but the ID certainly has no value on its own for strict formal reference or comparison. For that

reason, the Message-ID: SHOULD NOT be used for any function that has security implications.

[3.4.2.](#) ENVID

The ENVID (envelope identifier) can be used for message-tracking purposes [[RFC3885](#)] concerning a single posting/delivery transfer. The ENVID labels a single transit of the MHS by a specific message. So, the ENVID is used for one message posting, until that message is delivered. A re-posting of the message, such as by a Mediator, does not re-use that ENVID, but can use a new one, even though the message might legitimately retain its original Message-ID:.

The format of an ENVID is free form. Although its creator might choose to impose structure on the string, none is imposed by Internet standards. By implication, the scope of the string is defined by the domain name of the Return Address.

[4.](#) Services and Standards

The Internet Mail architecture comprises six basic types of

Crocker

Expires May 4, 2009

[Page 19]

Internet-Draft

EMail Architecture

October 2008

functionality, which are arranged to support a store-and-forward service. As shown in Figure 5, each type can have multiple instances, some of which represent specialized roles. This section considers the activities and relationships among these components, and the Internet Mail standards that apply to them.

Message

Mail User Agent (MUA)

Author MUA (aMUA)

Recipient MUA (rMUA)

Message Submission Agent (MSA)

Author-focused MSA functions (aMSA)

MHS-focused MSA functions (hMSA)

Message Transfer Agent (MTA)

Message Delivery Agent (MDA)

Recipient-focused MDA functions (rMDA)

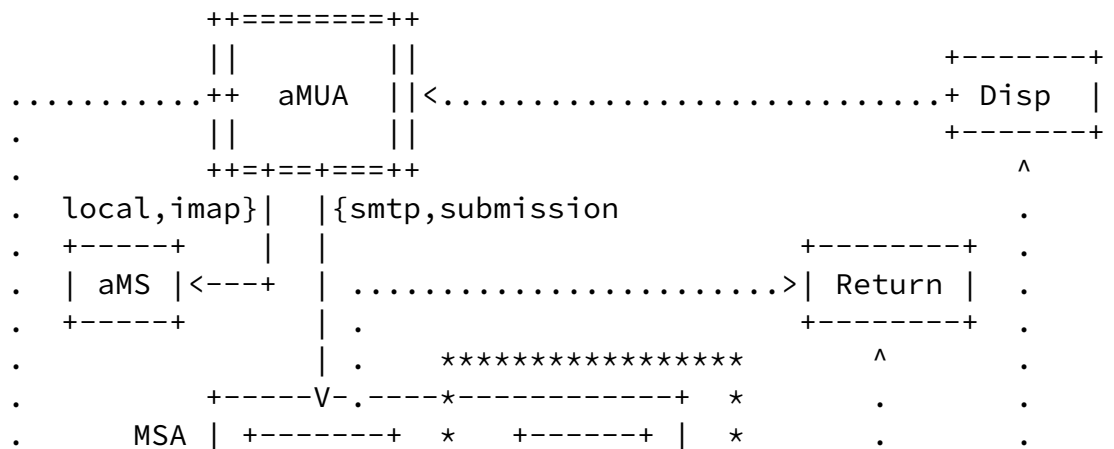
MHS-focused MDA functions (hMDA)

Message Store (MS)

Author MS (aMS)

Recipient MS (rMS)

This figure shows function modules and the standardized protocols used between them.



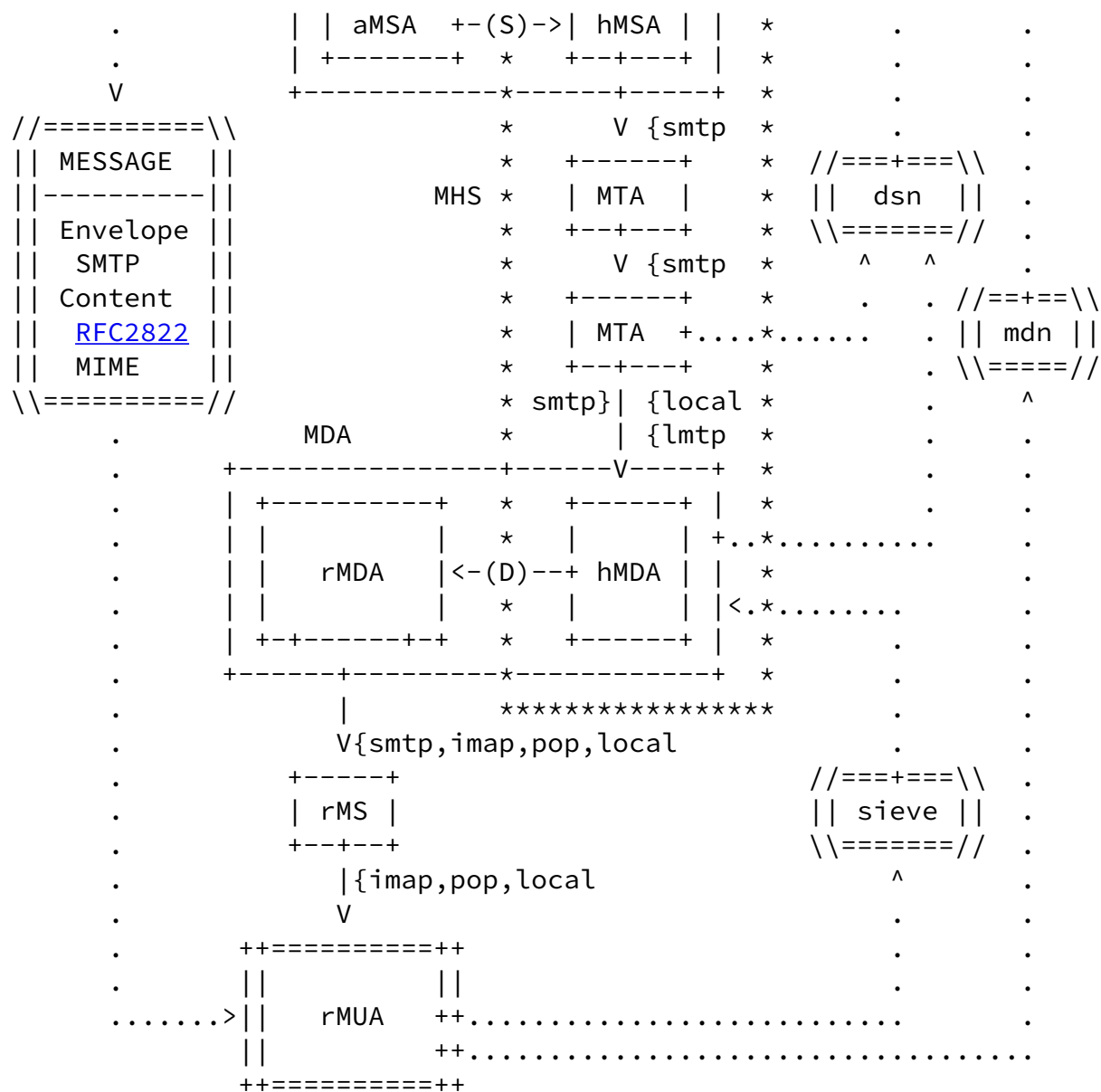


Figure 5: Protocols and Services

4.1. Message Data

The purpose of the Mail Handling Service (MHS) is to exchange a message object among participants [RFC2822], [RFC0822]. All of its underlying mechanisms serve to deliver that message from its Author

to its Recipients. A message can be explicitly labeled as to its nature [[RFC3458](#)].

A message comprises a transit-handling envelope and the message content. The envelope contains information used by the MHS. The content is divided into a structured header and the body. The header comprises transit handling trace information and structured fields that are part of the Author's message content. The body can be unstructured lines of text or a tree of multi-media subordinate objects, called "body-parts" or attachments [[RFC2045](#)], [[RFC2046](#)], [[RFC2047](#)], [[RFC4288](#)], [[RFC4289](#)], [[RFC2049](#)].

In addition, Internet Mail has a few conventions for special control data, notably:

Delivery Status Notification (DSN):

A Delivery Status Notification (DSN) is a message that can be generated by the MHS (MSA, MTA, or MDA) and sent to the [RFC2821](#).MailFrom address. An MDA and MTA are shown as sources of DSNs in Figure 5, and the destination is shown as Returns. DSNs provide information about message transit, such as transfer errors or successful delivery. [[RFC3461](#)]

Message Disposition Notification (MDN):

A Message Disposition Notification (MDN) is a message that provides information about post-delivery processing, such as indicating that the message has been displayed [[RFC3798](#)] or the form of content that can be supported [[RFC3297](#)]. It can be generated by an rMUA and is sent to the Disposition-Notification-To addresses. The mailbox for this is shown as Disp in Figure 5.

Message Filtering (SIEVE):

Sieve is a scripting language used to specify conditions for differential handling of mail, typically at the time of delivery [[RFC5228](#)]. Scripts can be conveyed in a variety of ways, as a MIME part. Figure 5 shows a Sieve script going from the rMUA to the MDA. However, filtering can be done at many different points along the transit path, and any one or more of them might be subject to Sieve directives, especially within a single ADMD. the Figure 5 shows only one relationship, for (relative) simplicity.

[4.1.1.](#) Envelope

Internet Mail has a fragmented framework for transit-related handling information. Information that is used directly by the MHS is called the "envelope." It directs handling activities by the transfer service and is carried in transfer service commands. That is, the envelope exists in the transfer protocol SMTP. [[RFC2821](#)]

Trace information, such as [RFC2822](#).Received, is recorded in the message header and is not subsequently altered. [[RFC2822](#)]

[4.1.2.](#) Header Fields

Header fields are attribute name/value pairs that cover an extensible range of email service parameters, structured user content, and user transaction meta-information. The core set of header fields is defined in [[RFC2822](#)], [[RFC0822](#)]. It is common practice to extend this set for different applications. Procedures for registering header fields are defined in [[RFC3864](#)]. An extensive set of existing header field registrations is provided in [[RFC4021](#)].

One danger of placing additional information in header fields is that Gateways often alter or delete them.

[4.1.3.](#) Body

The body of a message might be lines of ASCII text or a hierarchically structured composition of multi-media body-part attachments, using MIME. [[RFC2045](#)], [[RFC2046](#)], [[RFC2047](#)], [[RFC4288](#)], [[RFC2049](#)]

[4.1.4.](#) Identity References in a Message

Table 1 lists the core identifiers present in a message during transit.

Internet-Draft

EMail Architecture

October 2008

Layer	Field	Set By
Message Body	MIME Header	Author
Message header fields	From:	Author
	Sender:	Originator
	Reply-To:	Author
	To:, CC:, BCC:	Author
	Message-ID:	Originator
	Received:	Originator, Relay, Receiver
	Return-Path:	MDA, from MailFrom
	Resent-*:	Mediator
	List-Id:	Mediator
	List-*:	Mediator
SMTP	HELO/EHLO	Latest Relay Client
	ENVID	Originator
	MailFrom	Originator
	RcptTo	Author
	ORCPT	Author
IP	Source Address	Latest Relay Client

Table 1: Layered Identities

These are the most common address-related fields:

[RFC2822](#).From: Set by - Author

Names and addresses for authors of the message content are listed in the From: field.

[RFC2822](#).Reply-To: Set by - Author

If a Recipient sends a reply message that would otherwise use the [RFC2822](#).From field addresses in the original message, the addresses in the [RFC2822](#).Reply-To field are used instead. In other words, this field overrides the From: field for responses from Recipients.

[RFC2822](#).Sender: Set by - Originator

Crocker

Expires May 4, 2009

[Page 24]

Internet-Draft

EMail Architecture

October 2008

This field specifies the address responsible for submitting the message to the transfer service. This field can be omitted if it contains the same address as [RFC2822](#).From. However, omitting this field does not mean that no Sender is specified; it means that that header field is virtual and that the address in the From: field MUST be used.

Specification of the notifications Return addresses, which are contained in [RFC2821](#).MailFrom, is made by the [RFC2822](#).Sender. Typically the Return address is the same as the Sender address. However, some usage scenarios require it to be different.

[RFC2822](#).To/.CC: Set by - Author

These fields specify MUA Recipient addresses. However, some or all of the addresses in these fields might not be present in the [RFC2821](#).RcptTo commands.

The distinction between To and CC is subjective. Generally, a To addressee is considered primary and is expected to take action on the message. A CC addressee typically receives a copy as a courtesy.

[RFC2822](#).BCC: Set by - Author

A copy of the message might be sent to an addressee whose participation is not to be disclosed to the [RFC2822](#).To or [RFC2822](#).CC Recipients and, usually, not to the other BCC Recipients. The BCC: header field indicates a message copy to such a Recipient. Use of this field is discussed in [[RFC2822](#)].

[RFC2821](#).HELO/.EHLO: Set by - Originator, MSA, MTA

Any SMTP client -- including Originator, MSA, or MTA -- can specify its hosting domain identity for the SMTP HELO or EHLO

command operation.

[RFC3461](#).ENVID: Set by - Originator

The MSA can specify an opaque string, to be included in a DSN, as a means of assisting the Return address recipient in identifying the message that produced a DSN or message tracking.

[RFC2821](#).MailFrom: Set by - Originator

Crocker

Expires May 4, 2009

[Page 25]

Internet-Draft

E-Mail Architecture

October 2008

This field is an end-to-end string that specifies an email address for receiving return control information, such as returned messages. The name of this field is misleading, because it is not required to specify either the Author or the actor responsible for submitting the message. Rather, the actor responsible for submission specifies the [RFC2821](#).MailFrom address. Ultimately, the simple basis for deciding which address needs to be in the [RFC2821](#).MailFrom field is to determine which address must be informed about transfer-level problems (and possibly successes.)

[RFC2821](#).RcptTo: Set by - Author, Final MTA, MDA.

This field specifies the MUA mailbox address of a Recipient. The string might not be visible in the message content header. For example, the message destination address header fields, such as [RFC2822](#).To, might specify a mailing list mailbox, while the [RFC2821](#).RcptTo address specifies a member of that list.

[RFC2821](#).ORCPT: Set by - Author.

This is an optional parameter to the RCPT command, indicating the original address to which the current RCPT TO address corresponds, after a mapping was performed during transit. An ORCPT is the only reliable way to correlate a DSN from a multi-recipient message transfer with the intended recipient.

[RFC2821](#).Received: Set by - Originator, Relay, Mediator, Dest

This field contains trace information, including originating host, Relays, Mediators, and MSA host domain names and/or IP Addresses.

[RFC2821](#).Return-Path: Set by - Originator

The MDA records the [RFC2821](#).MailFrom address into the [RFC2822](#).Return-Path field.

[RFC2919](#).List-Id: Set by - Mediator Author

This field provides a globally unique mailing list naming framework that is independent of particular hosts. [[RFC2919](#)]

The identifier is in the form of a domain name; however, the string usually is constructed by combining the two parts of an email address. The result is rarely a true domain name, listed in the domain name service, although it can be.

[RFC2369](#).List-*: Set by - Mediator Author

[RFC2369] defines a collection of message header fields for use by mailing lists. In effect, they supply list-specific parameters for common mailing list user operations. The identifiers for these operations are for the list itself and the user-as-subscriber. [[RFC2369](#)]

[RFC0791](#).SourceAddr: Set by - The Client SMTP sending host immediately preceding the current receiving SMTP server.

[RFC0791] defines the basic unit of data transfer for the Internet: the IP Datagram. It contains a Source Address field that specifies the IP Address for the host (interface) from which the datagram was sent. This information is set and provided by the IP layer, which makes it independent of mail-level mechanisms. As such, it is often taken to be authoritative, although it is possible to provide false addresses.

[4.2](#). User-Level Services

Interactions at the user level entail protocol exchanges, distinct from those that occur at lower layers of the Internet Mail MHS architecture that is, in turn, above the Internet Transport layer. Because the motivation for email, and much of its use, is for interaction among people, the nature and details of these protocol exchanges often are determined by the needs of interpersonal and group communication. To accommodate the idiosyncratic behavior inherent in such communication, only subjective guidelines, rather than strict rules, can be offered for some aspects of system behavior. Mailing Lists provide particularly salient examples.

[4.2.1.](#) Mail User Agent (MUA)

A Mail User Agent (MUA) works on behalf of User actors and User applications. It is their representative within the email service.

The Author MUA (aMUA) creates a message and performs initial submission into the transfer infrastructure via a Mail Submission Agent (MSA). It can also perform any creation- and posting-time archival in its Message Store (aMS). An MUA aMS can organize messages in many different ways. A common model uses aggregations, called "folders". This model allows a folder for messages under development (Drafts), a folder for messages waiting to be sent (Queued or Unsent), and a folder for messages that have been successfully posted for transfer (Sent). But none of these folders is required. For example, IMAP allows drafts to be stored in any

folder; so no Drafts folder is present.

The Recipient MUA (rMUA) works on behalf of the Recipient to process received mail. This processing includes generating user-level disposition control messages, displaying and disposing of the received message, and closing or expanding the user communication loop by initiating replies and forwarding new messages.

NOTE: Although not shown in Figure 5, an MUA itself can have a distributed implementation, such as a "thin" user interface module on a constrained device such as a smartphone, with most of the MUA functionality running remotely on a more capable

server. An example of such an architecture might use IMAP [[RFC3501](#)] for most of the interactions between an MUA client and an MUA server. An approach for such scenarios is defined by [[RFC4550](#)].

A Mediator is special class of MUA. It performs message re-posting, as discussed in [Section 2.1](#).

An MUA can be automated, on behalf of a user who is not present at the time the MUA is active. One example is a bulk sending service that has a timed-initiation feature. These services are not to be confused with a mailing list Mediator, since there is no incoming message triggering the activity of the automated service.

A popular and problematic MUA is an automatic responder, such as one that sends out-of-office notices. This behavior might be confused with that of a Mediator, but this MUA is generating a new message. Automatic responders can annoy users of mailing lists unless they follow [[RFC3834](#)]. ***** The recommendations in [RFC 3834](#) are an important consequence of the addressing architecture of Internet Mail so they do help illustrate the architecture. *****

These identity fields are relevant to a typical MUA:

[RFC2822](#).From

[RFC2822](#).Reply-To

[RFC2822](#).Sender

[RFC2822](#).To, [RFC2822](#).CC

[RFC2822](#).BCC

[4.2.2](#). Message Store (MS)

An MUA can employ a long-term Message Store (MS). Figure 5 depicts

an Author's MS (aMS) and a Recipient's MS (rMS). An MS can be located on a remote server or on the same machine as the MUA.

An MS acquires messages from an MDA either by a local mechanism or by using POP or IMAP. The MUA accesses the MS either by a local mechanism or by using POP or IMAP. Using POP for message access, rather than bulk transfer, is rare, awkward, and largely non-standard.

[4.3.](#) MHS-Level Services

[4.3.1.](#) Mail Submission Agent (MSA)

A Mail Submission Agent (MSA) accepts the message submitted by the aMUA and enforces the policies of the hosting ADMD and the requirements of Internet standards. An MSA represents an unusual functional dichotomy. It represents the interests of the Author (aMUA) during message posting, to facilitate posting success; it also represents the interests of the MHS. In the architecture, these responsibilities are modeled, as shown in Figure 5, by dividing the MSA into two sub-components, aMSA and hMSA, respectively. Transfer of responsibility for a single message, from an Author's environment to the MHS, is called "posting". In Figure 5 it is marked as the (S) transition, within the MSA.

The hMSA takes transit responsibility for a message that conforms to the relevant Internet standards and to local site policies. It rejects messages that are not in conformance. The MSA performs final message preparation for submission and effects the transfer of responsibility to the MHS, via the hMSA. The amount of preparation depends upon the local implementations. Examples of oMSA tasks include adding header fields, such as Date: and Message-ID:, and modifying portions of the message from local notations to Internet standards, such as expanding an address to its formal [RFC2822](#) representation.

Historically, standards-based MUA/MSA message postings have used SMTP. [\[RFC2821\]](#) The standard currently preferred is SUBMISSION. [\[RFC4409\]](#) Although SUBMISSION derives from SMTP, it uses a separate TCP port and imposes distinct requirements, such as access authorization.

These identities are relevant to the MSA:

[RFC2821](#).HELO/.EHLO

[RFC3461](#).ENVID

[RFC2821](#).MailFrom

[RFC2821](#).RcptTo

[RFC2821](#).Received

[RFC0791](#).SourceAddr

[4.3.2](#). Mail Transfer Agent (MTA)

A Mail Transfer Agent (MTA) relays mail for one application-level "hop." It is like a packet-switch or IP router in that its job is to make routing assessments and to move the message closer to the Recipients. Of course, email objects are typically much larger than the payload of a packet or datagram, and the end-to-end latencies are typically much higher. Relaying is performed by a sequence of MTAs, until the message reaches a destination MDA. Hence, an MTA implements both client and server MTA functionality; it does not change addresses in the envelope or reformulate the editorial content. A change in data form, such as to MIME Content-Transfer-Encoding, is within the purview of an MTA, but removal or replacement of body content is not. An MTA also adds trace information.

[[RFC2505](#)]

NOTE: Within a destination ADMD, email relaying modules can make a variety of changes to the message, prior to delivery. In such cases, these modules are acting as Gateways, rather than MTAs.

Internet Mail uses SMTP [[RFC2821](#)], [[RFC0821](#)] primarily to effect point-to-point transfers between peer MTAs. Other transfer mechanisms include Batch SMTP [[RFC2442](#)] and ODMR [[RFC2645](#)]. As with most network layer mechanisms, the Internet Mail SMTP supports a basic level of reliability, by virtue of providing for retransmission after a temporary transfer failure. Unlike typical packet switches (and Instant Messaging services), Internet Mail MTAs are expected to store messages in a manner that allows recovery across service interruptions, such as host system shutdown. The degree of such

Internet-Draft

EMail Architecture

October 2008

robustness and persistence by an MTA can vary. The base SMTP specification provides a framework for protocol response codes. An extensible enhancement to this framework is defined in [[RFC5248](#)]

The primary routing mechanism for Internet Mail is the DNS MX record [[RFC1035](#)], which specifies an MTA through which the queried domain can be reached. This mechanism presumes a public, or at least a common, backbone that permits any attached MTA to connect to any other.

MTAs can perform any of these well-established roles:

Boundary MTA: An MTA that is part of an ADMD and interacts with MTAs in other ADMDs. This is also called a Border MTA. There can be different Boundary MTAs, according to the direction of mail-flow.

Outbound MTA: An MTA that relays messages to other ADMDs.

Inbound MTA: An MTA that receives inbound SMTP messages from MTA Relays in other ADMDs, for example, an MTA running on the host listed as the target of an MX record.

Final MTA: The MTA that transfers a message to the MDA.

These identities are relevant to the MTA:

[RFC2821](#).HELO/.EHLO

[RFC3461](#).ENVID

[RFC2821](#).MailFrom

[RFC2821](#).RcptTo

[RFC2822](#).Received: Set by - Relay Server

[RFC0791](#).SourceAddr

[4.3.3.](#) Mail Delivery Agent (MDA)

A transfer of responsibility from the MHS to a Recipient's environment (mailbox) is called "delivery." In the architecture, as depicted in Figure 5, delivery takes place within a Mail Delivery

Crocker

Expires May 4, 2009

[Page 31]

Internet-Draft

EMail Architecture

October 2008

Agent (MDA) and is shown as the (D) transition from the MHS-oriented MDA component (hMDA) to the Recipient-oriented MDA component (rMDA).

An MDA can provide distinctive, address-based functionality, made possible by its detailed information about the properties of the destination address. This information might also be present elsewhere in the Recipient's ADMD, such as at an organizational border (Boundary) Relay. However, it is required for the MDA, if only because the MDA is required to know where to deliver the message.

Like an MSA, an MDA serves two roles, as depicted in Figure 5. Formal transfer of responsibility, called "delivery", is effected between the two components that embody these roles as shown as "(D)" in Figure 5. The MHS portion (hMDA) primarily functions as a server SMTP engine. A common additional role is to re-direct the message to an alternative address, as specified by the recipient addressee's preferences. The job of the recipient portion of the MDA (rMDA) is to perform any delivery actions that the Recipient specifies.

Transfer into the MDA is accomplished by a normal MTA transfer mechanism. Transfer from an MDA to an MS uses an access protocol, such as POP or IMAP.

NOTE: The term "delivery" can refer to the formal, MHS function specified here or to the first time a message is displayed to a Recipient. A simple, practical test for whether the MHS-based definition applies is whether a DSN can be generated.

These identities are relevant to the MDA:

[RFC2821](#).Return-Path: Set by - Author Originator or Mediator

Originator

The MDA records the [RFC2821](#).MailFrom address into the [RFC2822](#).Return-Path field.

[RFC2822](#).Received: Set by - MDA server

An MDA can record a Received: header field to indicate trace information, including source host and receiving host domain names and/or IP Addresses.

Crocker

Expires May 4, 2009

[Page 32]

Internet-Draft

EMail Architecture

October 2008

[4.4.](#) Transition Modes

From the origination site to the point of delivery, Internet Mail usually follows a "push" model. That is, the actor that holds the message initiates transfer to the next venue, typically with SMTP [[RFC2821](#)] or LMTP [[RFC2033](#)]. With a "pull" model, the actor that holds the message waits for the actor in the next venue to initiate a request for transfer. Standardized mechanisms for pull-based MHS transfer are ETRN [[RFC1985](#)] and ODMR [[RFC2645](#)].

After delivery, the Recipient's MUA (or MS) can gain access by having the message pushed to it or by having the receiver of access pull the message, such as by using POP [[RFC1939](#)] and IMAP [[RFC3501](#)].

[4.5.](#) Implementation and Operation

A discussion of any interesting system architecture often bogs down when architecture and implementation are confused. An architecture defines the conceptual functions of a service, divided into discrete conceptual modules. An implementation of that architecture can combine or separate architectural components, as needed for a particular operational environment. For example, a software system that primarily performs message relaying is an MTA, yet it might also include MDA functionality. That same MTA system might be able to interface with non-Internet email services and thus perform both as an MTA and as a Gateway.

Similarly, implemented modules might be configured to form elaborations of the architecture. An interesting example is a

distributed MS. One portion might be a remote server and another might be local to the MUA. As discussed in [[RFC1733](#)], there are three operational relationships among such MSs:

Online: The MS is remote, and messages are accessible only when the MUA is attached to the MS so that the MUA will re-fetch all or part of a message, from one session to the next.

Offline: The MS is local to the user, and messages are completely moved from any remote store, rather than (also) being retained there.

Disconnected: An rMS and a uMS are kept synchronized, for all or part of their contents, while they are connected. When they are disconnected, mail can arrive at the rMS and the user can make changes to the uMS. The two stores are re-synchronized when they are reconnected.

Crocker

Expires May 4, 2009

[Page 33]

Internet-Draft

EMail Architecture

October 2008

[5](#). Mediators

Basic message transfer from Author to Recipients is accomplished by using an asynchronous store-and-forward communication infrastructure in a sequence of independent transmissions through some number of MTAs. A very different task is a sequence of postings and deliveries through Mediators. A Mediator forwards a message, through a re-posting process. The Mediator shares some functionality with basic MTA relaying, but has greater flexibility in both addressing and content than is available to MTAs.

This is the core set of message information that is commonly set by all types of Mediators:

[RFC2821](#).HELO/.EHL0: Set by - Mediator Originator

[RFC3461](#).ENVID: Set by - Mediator Originator

[RFC2821](#).RcptTo: Set by - Mediator Author

The Mediator can record received information, to indicate the delivery to the original address and submission to the alias address. The trace of Received: header fields can include everything from original posting, through relaying, to final delivery.

The aspect of a Mediator that distinguishes it from any other MUA creating a message is that a Mediator preserves the integrity and tone of the original message, including the essential aspects of its origination information. The Mediator might also add commentary.

Examples of MUA messages that a Mediator does not create include:

New message that forwards an existing message:

Although this action provides a basic template for a class of Mediators, its typical occurrence is not, itself, an example of a Mediator. The new message is viewed as being from the actor that is doing the forwarding, rather than from the original Author.

A new message encapsulates the original message and is seen as from the new Originator. This Mediator Originator might add commentary and can modify the original message content.

Because the forwarded message is a component of the message sent by the new Originator, the new message creates a new dialogue. However the final Recipient still sees the contained message as from the original Author.

Reply:

When a Recipient responds to the Author of a message, the new message is not typically viewed as a forwarding of the original. Its focus is the new content, although it might contain all or part of the material from the original message. The earlier material is merely contextual and secondary. This includes automated replies, such as vacation out-of-office notices, as discussed in [Section 4.2.1](#).

Annotation:

The integrity of the original message is usually preserved, but one or more comments about the message are added in a manner that distinguishes commentary from original text. The primary purpose of the new message is to provide commentary from a new Author, similar to a Reply.

The remainder of this section describes common examples of Mediators.

[5.1.](#) Alias

One function of an MDA is to determine the internal location of a mailbox in order to perform delivery. An Alias is a simple re-addressing facility that provides one or more new Internet Mail addresses, rather than a single, internal one; the message continues through the transfer service, for delivery to one or more alternate addresses. Although typically implemented as part of an MDA, this facility is a Recipient function. It resubmits the message, although all handling information except the envelope recipient ([rfc2821](#).RcptTo) address is retained. In particular, the Return address ([rfc2821](#).MailFrom) is unchanged.

What is distinctive about this forwarding mechanism is how closely it resembles normal MTA store-and-forward relaying. Its only significant difference is that it changes the [RFC2821](#).RcptTo value. Because this change is so small, aliasing can be viewed as a part of the lower-level mail relaying activity. However, this small change has a large semantic impact: The designated recipient has chosen a new recipient.

NOTE: When the replacement list includes more than one address, the alias is increasingly likely to have delivery problems. Any problem reports go to the original Author, not the administrator of the alias entry. This makes it more difficult to resolve the problem, because the original Author has no knowledge of the Alias mechanism.

Alias typically changes only envelope information:

[RFC2822](#).To/.CC/.BCC: Set by - Author

These fields retain their original addresses.

[RFC2821](#).MailFrom: Set by - Author

The benefit of retaining the original MailFrom value is to ensure that an actor related to the originating ADMD knows there has been a delivery problem. On the other hand, the responsibility for handling problems, when transiting from the original recipient mailbox to the alias mailbox usually lies with that original Recipient, because the Alias mechanism is strictly under that Recipient's control. Retaining the original MailFrom address prevents this.

[5.2](#). ReSender

Also called the ReDirector, the ReSender's actions differ from forwarding because the Mediator "splices" a message's addressing information to connect the Author of the original message with the Recipient of the new message. This connection permits them to have direct exchange, using their normal MUA Reply functions, while also recording full reference information about the Recipient who served as a Mediator. Hence, the new Recipient sees the message as being from the original Author, even if the Mediator adds commentary.

These identities are relevant to a resent message:

[RFC2822](#).From: Set by - original Author

content are retained. The free-form (display-name) portion of the address might be modified to provide informal reference to the ReSender.

[RFC2822](#).Reply-To: Set by - original Author

If this field is present in the original message, it is retained in the resent message.

[RFC2822](#).Sender: Set by - Author's Originator or Mediator Originator.

[RFC2822](#).To/.CC/.BCC: Set by - original Author

These fields specify the original message Recipients.

[RFC2822](#).Resent-From: Set by - Mediator Author

This address is of the original Recipient who is redirecting the message. Otherwise, the same rules apply to the Resent-From: field as to an original [RFC2822](#).From field.

[RFC2822](#).Resent-Sender: Set by - Mediator Originator

The address of the actor responsible for resubmitting the message. As with [RFC2822](#).Sender, this field can be omitted when it contains the same address as [RFC2822](#).Resent-From.

[RFC2822](#).Resent-To/-CC/-BCC: Set by: Mediator Author

The addresses of the new Recipients who are now able to reply to the original author.

[RFC2821](#).MailFrom: Set by - Mediator Originator

The actor responsible for resubmission ([RFC2822](#).Resent-Sender) is also responsible for specifying the new MailFrom address.

[5.3](#). Mailing Lists

A Mailing List receives messages as an explicit addressee and then re-posts them to a list of subscribed members. The Mailing List performs a task that can be viewed as an elaboration of the ReSender. In addition to sending the new message to a potentially large number of new Recipients, the Mailing List can modify content, for example, by deleting attachments, converting the format, and adding list-specific comments. Mailing Lists also archive messages posted by

Authors. Still the message retains characteristics of being from the original Author.

These identities are relevant to a mailing list processor, when submitting a message:

[RFC2919](#).List-Id: Set by - Mediator Author

[RFC2369](#).List-*: Set by - Mediator Author

[RFC2822](#).From: Set by - original Author

Names and email addresses for the original Author of the message content are retained.

[RFC2822](#).Reply-To: Set by - Mediator or original Author

Although problematic, it is common for a Mailing List to assign its own addresses to the Reply-To: header field of messages that it posts. This assignment is intended to ensure that replies go to all list members, rather than to only the original Author. As a User actor, a Mailing List is the Author of the new message and can legitimately set the Reply-To: value. As a Mediator attempting to represent the message on behalf of its original Author, creating or modifying a Reply-To: field can be viewed as violating that Author's intent. Modifying the field to include the list address can send to the entire list replies that are meant only for the original Author. When the Mailing List does not set the field, a reply meant for the entire list can instead go only to the original Author. At best, either choice is a matter of group culture for the particular list.

[RFC2822](#).Sender: Set by - Author Originator or Mediator Originator

This field usually specifies the address of the actor responsible for Mailing List operations. Mailing Lists that operate in a manner similar to a simple MTA Relay preserve as much of the original handling information as possible, including the original [RFC2822](#).Sender field. (Note that this mode of operation causes the Mailing List to behave much like an Alias, with a possible difference in number of new addressees.)

Internet-Draft

EMail Architecture

October 2008

[RFC2822](#).To/.CC: Set by - original Author

These fields usually contain the original list of Recipient addresses.

[RFC2821](#).MailFrom: Set by - Mediator Originator

Because a Mailing List can modify the content of a message in any way, it is responsible for that content; that is, it is an Author. As such, the Return Address is specified by the Mailing List. Although it is plausible for the Mailing List to re-use the Return Address employed by the original Originator, notifications sent to that address after a message has been processed by a Mailing List could be problematic.

[5.4](#). Gateways

A Gateway performs the basic routing and transfer work of message relaying, but it also is permitted to modify content, structure, address, or attributes as needed to send the message into a messaging environment that operates under different standards or potentially incompatible policies. When a Gateway connects two differing messaging services, its role is easy to identify and understand. When it connects environments that follow similar technical standards, but significantly different administrative policies, it is easy to view a Gateway as merely an MTA.

The critical distinction between an MTA and a Gateway is that a Gateway can make substantive changes to a message to map between the standards. In virtually all cases, this mapping results in some degree of semantic loss. The challenge of Gateway design is to minimize this loss. Standardized gateways to Internet Mail are facsimile [[RFC4143](#)], voicemail [[RFC3801](#)], and MMS [[RFC4356](#)]

A Gateway can set any identity field available to an MUA. These identities are typically relevant to Gateways:

[RFC2822](#).From: Set by - original Author

Names and addresses for the original Author of the message content are retained. As for all original addressing information in the message, the Gateway can translate addresses as required to continue to be useful in the target environment.

Crocker

Expires May 4, 2009

[Page 39]

Internet-Draft

EMail Architecture

October 2008

[RFC2822](#).Reply-To: Set by - original Author

The Gateway SHOULD retain this information, if it is present. The ability to perform a successful reply by a Recipient is a typical test of Gateway functionality.

[RFC2822](#).Sender: Set by - Author Originator or Mediator
Originator

This field can retain the original value or can be set to a new address.

[RFC2822](#).To/.CC/.BCC: Set by - original Recipient

These fields usually retain their original addresses.

[RFC2821](#).MailFrom: Set by - Author Originator or Mediator
Originator

The actor responsible for handling the message can specify a new address to receive handling notices.

[5.5](#). Boundary Filter

To enforce security boundaries, organizations can subject messages to analysis, for conformance with its safety policies. An example is detection of content classed as spam or a virus. A filter might alter the content, to render it safe, such as by removing content deemed unacceptable. Typically, these actions add content to the message that records the actions.

[6.](#) Considerations

[6.1.](#) Security Considerations

This document describes the existing Internet Mail architecture. It introduces no new capabilities. The security considerations of this deployed architecture are documented extensively in the technical specifications referenced by this document. These specifications cover classic security topics, such as authentication and privacy. For example, email transfer protocols can use standardized mechanisms for operation over authenticated and/or encrypted links, and message content has similar protection standards available. Examples of such mechanisms include SMTP-TLS [[RFC3207](#)], SMTP-Auth [[RFC2554](#)], OpenPGP [[RFC4880](#)], and S/MIME [[RFC3851](#)].

The core of the Internet Mail architecture does not impose any

security requirements or functions on the end-to-end or hop-by-hop components. For example, it does not require participant authentication and does not attempt to prevent data disclosure.

Particular message attributes might expose specific security considerations. For example, the blind carbon copy feature of the architecture invites disclosure concerns, as discussed in [section 7.2 of \[RFC2821\]](#) and [section 5 of \[RFC2822\]](#). Transport of text or non-text content in this architecture has security considerations that are discussed in [[RFC2822](#)], [[RFC2045](#)], [[RFC2046](#)], and [[RFC4288](#)] as well as the security considerations present in the IANA media types registry for the respective types.

Agents that automatically respond to email raise significant security considerations, as discussed in [[RFC3834](#)]. Gateway behaviors affect end-to-end security services, as discussed in [[RFC2480](#)]. Security considerations for boundary filters are discussed in [[RFC5228](#)].

See [section 7.1 of \[RFC2821\]](#) for a discussion of the topic of origination validation. As mentioned in [Section 4.1.4](#), it is common practice for components of this architecture to use the [[RFC0791](#)].SourceAddr to make policy decisions [[RFC2505](#)], although the address can be "spoofed". It is possible to use it without authorization. SMTP and Submission authentication [[RFC2554](#)], [[RFC4409](#)] provide more secure alternatives.

The discussion of trust boundaries, ADMDs, actors, roles, and responsibilities in this document highlights the relevance and potential complexity of security factors for operation of an Internet mail service. The core design of Internet Mail to encourage open and casual exchange of messages has met with scaling challenges, as the population of email participants has grown to include those with problematic practices. For example, spam, as defined in [\[RFC2505\]](#), is a by-product of this architecture. A number of standards track or BCP documents on the subject have been issued. [\[RFC2505\]](#), [\[RFC5068\]](#), [\[RFC3685\]](#)

[6.2.](#) IANA Considerations

This document has no actions for IANA.

[6.3.](#) Internationalization

Because its origins date back to the use of ASCII, Internet Mail has had an ongoing challenge to support the wide range of necessary international data representations. For a discussion of this topic, see [\[MAIL-I18N\]](#).

[7.](#) References

[7.1.](#) Normative

- [RFC0791] Postel, J., "Internet Protocol", [RFC 791](#), 1981 September.
- [RFC1034] Mockapetris, P., "Domain names - concepts and facilities", STD 13, [RFC 1034](#), November 1987.
- [RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, [RFC 1035](#), November 1987.
- [RFC1939] Myers, J. and M. Rose, "Post Office Protocol - Version 3", STD 53, [RFC 1939](#), May 1996.
- [RFC2045] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies", [RFC 2045](#), November 1996.

- [RFC2046] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types", [RFC 2046](#), November 1996.
- [RFC2047] Moore, K., "MIME (Multipurpose Internet Mail Extensions) Part Three: Message Header Extensions for Non-ASCII Text", [RFC 2047](#), November 1996.
- [RFC2049] Freed, N. and N. Borenstein, "Multipurpose Internet Mail Extensions (MIME) Part Five: Conformance Criteria and Examples", [RFC 2049](#), November 1996.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC2181] Elz, R. and R. Bush, "Clarifications to the DNS Specification", [RFC 2181](#), July 1997.
- [RFC2369] Neufeld, G. and J. Baer, "The Use of URLs as Meta-Syntax for Core Mail List Commands and their Transport through Message Header Fields", [RFC 2369](#), July 1998.
- [RFC2645] "On-Demand Mail Relay (ODMR) SMTP with Dynamic IP Addresses", [RFC 2645](#), August 1999.
- [RFC2821] Klensin, J., "Simple Mail Transfer Protocol", [RFC 2821](#), April 2001.
- [RFC2822] Resnick, P., "Internet Message Format", [RFC 2822](#),

April 2001.

- [RFC2919] Chandhok, R. and G. Wenger, "List-Id: A Structured Field and Namespace for the Identification of Mailing Lists", [RFC 2919](#), March 2001.
- [RFC3192] Allocchio, C., "Minimal FAX address format in Internet Mail", [RFC 2304](#), October 2001.
- [RFC3297] Klyne, G., Iwazaki, R., and D. Crocker, "Content Negotiation for Messaging Services based on Email",

[RFC 3297](#), July 2002.

- [RFC3458] Burger, E., Candell, E., Eliot, C., and G. Klyne, "Message Context for Internet Mail", [RFC 3458](#), January 2003.
- [RFC3461] Moore, K., "Simple Mail Transfer Protocol (SMTP) Service Extension for Delivery Status Notifications (DSNs)", [RFC 3461](#), January 2003.
- [RFC3501] Crispin, M., "Internet Message Access Protocol - Version 4rev1", [RFC 3501](#), March 2003.
- [RFC3798] Hansen, T. and G. Vaudreuil, "Message Disposition Notification", [RFC 3798](#), May 2004.
- [RFC3834] Moore, K., "Recommendations for Automatic Responses to Electronic Mail", [RFC 3834](#), August 2004.
- [RFC3864] Klyne, G., Nottingham, M., and J. Mogul, "Registration Procedures for Message Header Fields", [RFC 3864](#), September 2004.
- [RFC4021] Klyne, G. and J. Palme, "Registration of Mail and MIME Header Fields", [RFC 4021](#), March 2005.
- [RFC4288] Freed, N., Klensin, J., and J. Postel, "Media Type Specifications and Registration Procedures", [BCP 13](#), [RFC 4288](#), December 2005.
- [RFC4289] Freed, N., Klensin, J., and J. Postel, "Multipurpose Internet Mail Extensions (MIME) Part Four: Registration Procedures", [BCP 13](#), [RFC 4289](#), December 2005.
- [RFC4409] Gellens, R. and J. Klensin, "Message Submission for Mail", [RFC 4409](#), April 2006.
- [RFC4550] Maes, S., , S., and Isode Ltd., "Internet Email to Support

Diverse Service Environments (Lemonade) Profile",
June 2006.

- [RFC5228] Showalter, T., "Sieve: A Mail Filtering Language",

[RFC 5228](#).

[RFC5248] Hansen, T. and J. Klensin, "A Registry for SMTP Enhanced Mail System Status Codes", [RFC 5248](#), June 2008.

[7.2](#). Informative

[MAIL-I18N]

Internet Mail Consortium, "Using International Characters in Internet Mail", IMC IMCR-010, August 1998.

[RFC0821] Postel, J., "Simple Mail Transfer Protocol", STD 10, [RFC 821](#), August 1982.

[RFC0822] Crocker, D., "Standard for the format of ARPA Internet text messages", STD 11, [RFC 822](#), August 1982.

[RFC1733] Crispin, M., "Distributed Electronic Models in IMAP4", December 1994.

[RFC1767] Crocker, D., "MIME Encapsulation of EDI Objects", [RFC 1767](#), March 1995.

[RFC1985] De Winter, J., "SMTP Service Extension for Remote Message Queue Starting", August 1996.

[RFC2033] Myers, J., "Local Mail Transfer Protocol", [RFC 2033](#), October 1996.

[RFC2142] Crocker, D., "Mailbox Names for Common services, Roles and Functions", [RFC 2142](#), May 1997.

[RFC2442] "The Batch SMTP Media Type", [RFC 2442](#), November 1998.

[RFC2480] Freed, N., "Gateways and MIME Security Multiparts", [RFC 2480](#), January 1999.

[RFC2505] Lindberg, G., "Anti-Spam Recommendations for SMTP MTAs", [RFC 2505](#), February 1999.

[RFC2554] Myers, J., "SMTP Service Extension for Authentication", [RFC 2554](#), March 1999.

[RFC3207] Hoffman, P., "SMTP Service Extension for Secure SMTP over

Transport Layer Security", [RFC 3207](#), February 2002.

- [RFC3685] Daboo, C., "SIEVE Email Filtering: Spamtest and VirusTest Extensions", [RFC 3685](#), February 2004.
- [RFC3801] Vaudreuil, G. and G. Parsons, "Voice Profile for Internet Mail - version 2 (VPIMv2)", [RFC 3801](#), June 2004.
- [RFC3851] Ramsdell, B., Ed., "Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.1 Message Specification", [RFC 3851](#), July 2004.
- [RFC3885] Allman, E. and T. Hansen, "SMTP Service Extension for Message Tracking", [RFC 3885](#), September 2004.
- [RFC4142] Crocker, D. and G. Klyne, "Full-mode Fax Profile for Internet Mail: FFPIM", December 2005.
- [RFC4143] Toyoda, K. and D. Crocker, "Facsimile Using Internet Mail (IFAX) Service of ENUM", [RFC 4143](#), November 2005.
- [RFC4356] Gellens, R., "Mapping Between the Multimedia Messaging Service (MMS) and Internet Mail", [RFC 4356](#), January 2006.
- [RFC4880] Callas, J., Donnerhacke, L., Finney, H., Shaw, D., and R. Thayer, "OpenPGP Message Format", [RFC 4880](#), November 2007.
- [RFC5068] Hutzler, C., Crocker, D., Resnick, P., Sanderson, R., and E. Allman, "Email Submission Operations: Access and Accountability Requirements", [RFC 5068](#), [BCP 134](#), Nov 2007.
- [Tussle] Clark, D., Wroclawski, J., Sollins, K., and R. Braden, "Tussle in Cyberspace: Defining Tomorrow's Internet", ACM SIGCOMM, 2002.

[Appendix A](#). Acknowledgements

This work derives from a section in an early version of [\[RFC5068\]](#). Discussion of the Originator actor role was greatly clarified during discussions in the IETF's Marid working group.

Graham Klyne, Pete Resnick and Steve Atkins provided thoughtful insight on the framework and details of the original drafts, as did Chris Newman for the final versions, while also serving as cognizant Area Director for the document. Tony Hansen served as document shepherd, through the IETF process.

Internet-Draft

EMail Architecture

October 2008

Later reviews and suggestions were provided by Eric Allman, Nathaniel Borenstein, Ed Bradford, Cyrus Daboo, Frank Ellermann, Tony Finch, Ned Freed, Eric Hall, Willemien Hoogendoorn, Brad Knowles, John Leslie, Bruce Valdis Kletnieks, Mark E. Mallett, David MacQuigg, Alexey Melnikov, der Mouse, S. Moonesamy, Daryl Odnert, Rahmat M. Samik-Ibrahim, Marshall Rose, Hector Santos, Jochen Topf, Greg Vaudreuil.

Diligent early proof-reading was performed by Bruce Lilly. Diligent technical editing was provided by Susan Hunziker

Index

10

A

- accountability 11
- accountable 12-13
- Actor
 - Administrative 13
 - Author 8
 - Consumer 14
 - Edge 14
 - Gateway 12
 - Originator 11
 - Recipient 9
 - Return Handler 9
 - Transit 14
- Actors
 - MHS 10
- ADMD 11, 13-14, 18, 23, 29, 36
- Administrative Actors 13
- Administrative Management Domain 11
- aMSA 29
- Author 8, 10
- author 33

B

- body-parts 22
- bounce handler 9

boundary 14

C

Consumer Actor 14

content 10, 12-13, 18, 22, 30

D

delivery 4, 9-10, 12-13, 17, 22-23, 33, 35-36

Crocker

Expires May 4, 2009

[Page 46]

Internet-Draft

EMail Architecture

October 2008

Discussion of document 7

E

Edge Actor 14

end-to-end 4

envelope 9, 12, 19, 22-23, 30, 35-36

ETRN 33

G

Gateway 10, 12

H

header 22

hMSA 29

I

Internet Mail 4

L

LMTP 33

local-part 16

M

Mail 4

Mail User Agent 4

Mail From 35

Mail Handling Service 4, 10

Mail Submission Agent 11

Mail Transfer Agent 4

mailbox 35

MDA 35

MDN 9

message 6, 22

Message Disposition Notification 9

MHS 4, 9-12, 19-20, 22-23

Actors 10

MSA 11, 29

MTA 4, 14

boundary 14

MUA 4, 13, 28-29

O

ODMR 33

Originator 10-11

P

posting 4, 9, 11, 19, 28-29, 33, 36

pull 33

Crocker

Expires May 4, 2009

[Page 47]

Internet-Draft

EMail Architecture

October 2008

push 33

R

RcptTo 10

Receiver 10

Recipient 9-10, 35

recipient 33

relay 10

responsibility 29

responsible 12-13

Return address 35

Return Handler 9

role 9, 17

Author 8

Originator 11

Recipient 9

S

SIEVE 22

SMTP 33

T

transfer 10, 12-13

Transit Actor 14

transition 29

U

UA 4

User Agent 4

Author's Address

Dave Crocker
Brandenburg InternetWorking
675 Spruce Drive
Sunnyvale, CA 94086
USA

Phone: +1.408.246.8253

Email: dcrocker@bbiw.net

Crocker

Expires May 4, 2009

[Page 48]

Internet-Draft

E-Mail Architecture

October 2008

Full Copyright Statement

Copyright (C) The IETF Trust (2008).

This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.